Molecular Informatics 2018

Predictive Models for Kinetic Parameters of Cycloaddition Reactions

Kazan Federal University, 420008, Kremlevskaya 18, Kazan, Russia

Abstract

© 2018 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim This paper reports SVR (Support Vector Regression) and GTM (Generative Topographic Mapping) modeling of three kinetic properties of cycloaddition reactions: rate constant (logk), activation energy (Ea) and pre-exponential factor (logA). A data set of 1849 reactions, comprising (4+2), (3+2) and (2+2) cycloadditions (CA) were studied in different solvents and at different temperatures. The reactions were encoded by the ISIDA fragment descriptors generated for Condensed Graph of Reaction (CGR). For a given reaction, a CGR condenses structures of all the reactants and products into one single molecular graph, described both by conventional chemical bonds and “dynamical” bonds characterizing chemical transformations. Different scenarios of logk assessment were exploited: direct modeling, application of the Arrhenius equation and temperature-scaled GTM landscapes. The logk models with optimal cross-validated statistics (Q2=0.78–0.94 RMSE=0.45–0.86) have been challenged to predict rates for the external test set of 200 reactions, comprising both reactions that were not present in the training set, and training set transformations performed under different reaction conditions. The models are freely available on our web-server: http://cimm.kpfu.ru/models.

http://dx.doi.org/10.1002/minf.201800077

Keywords

Condensed Graph of Reaction, cycloaddition reactions, Generative Topographic Mapping, QSPR

References

[1] K. N. Houk, Acc. Chem. Res. 1975, 8, 361–369
[2] S. Kahn, C. Pau, L. Overman, W. J. Hehre, J. Am. Chem. Soc. 1986, 108, 7381–7396
[3] J. Chandrasekhar, S. Shariffskul, W. L. Jorgensen, J. Phys. Chem. B 2002, 106, 8078–8085
[4] O. Acevedo, W. L. Jorgensen, J. Chem. Theory Comput. 2007, 3, 1412–1419
[5] O. Acevedo, W. L. Jorgensen, J. D. Evanseck, J. Chem. Theory Comput. 2007, 3, 132–138
[6] V. D. Kiselev, A. N. Ustiugov, I. P. Breus, A. I. Konovalov, Dokl. Akad. Nauk SSSR 1977, 234, 1089–1092
[7] V. D. Kiselev, A. I. Konovalov, J. Phys. Org. Chem. 2009, 22, 466–483
[8] G. Szabo, L. Nyulaszi, Struct. Chem. 2017, 28, 333–343
[9] X. Long, J. Niu, Chemosphere 2007, 67, 2028–2034
[10] T. C. Ho, Catal. Rev. 2008, 50, 287–378
[11] A. Fernández-Ramos, J. A. Miller, S. J. Klippenstein, D. G. Truhlar, Chem. Rev. 2006, 106, 4518–4584
[12] A. A. Toropov, V. O. Kudyshkin, N. L. Voropaeva, I. N. Ruban, S. S. Rashidova, J. Struct. Chem. 2004, 45, 945–950

[13] X. Yu, B. Yi, X. Wang, Eur. Polym. J. 2008, 44, 3997–4001
[14] J. Gasteiger, U. Hondelmann, P. Röse, W. Witzenbichler, J. Chem. Soc. Perkin Trans. 2 1995, 193–204
[15] X. H. Jin, S. Peldszus, P. M. Huck, Chemosphere 2015, 138, 1–9
[16] D. A. Latino, J. Aires-de-Sousa, J. Chem. Inf. Model. 2009, 49, 1839–1846
[17] J. A. Morrill, J. H. Biggs, C. N. Bowman, J. W. Stansbury, J. Mol. Graphics Modell. 2011, 29, 763–772
[18] W. A. Warr, Mol. Inf. 2014, 33, 469–476
[19] I. I. Baskin, T. I. Madzhidov, I. S. Antipin, A. A. Varnek, Russ. Chem. Rev. 2017, 86, 1127
[20] Y. E. Zevatskii, D. V. Samoilov, Russ. J. Org. Chem. 2007, 43, 483–500
[21] G. Marcou, J. Aires de Sousa, D. A. Latino, A. de Luca, D. Horvath, V. Rietsch, A. Varnek, J. Chem. Inf. Model. 2015, 55, 239–250
[22] A. Kravtsov, P. Karpov, I. Baskin, V. Palyulin, N. Zefirov, in Doklady Chemistry, Vol. 440, Springer, 2011, pp. 299–301
[23] T. Madzhidov, P. Polishchuk, R. Nugmanov, A. Bodrov, A. Lin, I. Baskin, A. Varnek, I. Antipin, Russ. J. Org. Chem. 2014, 50, 459–463
[24] T. I. Madzhidov, A. V. Bodrov, T. R. Gimadiev, R. I. Nugmanov, I. S. Antipin, A. A. Varnek, J. Struct. Chem. 2015, 56, 1227–1234
[25] P. Polishchuk, T. Madzhidov, T. Gimadiev, A. Bodrov, R. Nugmanov, A. Varnek, J. Comput.-Aided Mol. Des. 2017, 31, 829–839
[26] F. Hoonakker, N. Lachiche, A. Varnek, A. Wagner, Int. J. Artif. Intell. Tools 2011, 20, 253–270
[27] T. I. Madzhidov, P. G. Polishchuk, R. I. Nugmanov, A. V. Bodrov, A. I. Lin, I. I. Baskin, A. A. Varnek, I. S. Antipin, Russ. J. Org. Chem. 2014, 50, 459–463
[28] T. R. Gimadiev, T. I. Madzhidov, R. I. Nugmanov, I. I. Baskin, I. S. Antipin, A. Varnek, J. Comput.-Aided Mol. Des. 2018
[29] H. Drucker, C. J. Burges, L. Kaufman, A. J. Smola, V. Vapnik, in Advances in neural information processing systems, 1997, pp. 155–161
[30] C. M. Bishop, M. Svensen, C. K. I. Williams, Neural Comput. 1998, 10, 215–234
[31] H. A. Gaspar, G. Marcou, D. Horvath, A. Arault, S. Lozano, P. Vayer, A. Varnek, J. Chem. Inf. Model. 2013, 53, 3318–3325
[32] N. Kireeva, S. L. Kuznetsov, A. Y. Tsivadze, Ind. Eng. Chem. Res. 2012, 51, 14337–14343
[33] H. A. Gaspar, Baskin, II, G. Marcou, D. Horvath, A. Varnek, Mol. Inf. 2015, 34, 348–356
[34] P. Sidorov, H. Gaspar, G. Marcou, A. Varnek, D. Horvath, J. Comput.-Aided Mol. Des. 2015, 29, 1087–1108
[35] T. R. Gimadiev, T. I. Madzhidov, G. Marcou, A. Varnek, BioNanoScience 2016, 6, 464–472
[36] Standardizer, 6.1.5, ChemAxon (http://www.chemaxon.com), 2013
[37] J. Catalán, V. López, P. Pérez, R. Martin-Villamil, J.-G. Rodríguez, Liebigs Ann. 1995, 1995, 241–252
[38] J. Catalán, C. Díaz, Liebigs Ann. 1997, 1997, 1941–1949
[39] J. Catalán, C. Díaz, V. López, P. Pérez, J.-L. G. De Paz, J. G. Rodríguez, Liebigs Ann. 1996, 1996, 1785–1794
[40] R. W. Taft, M. J. Kamlet, J. Am. Chem. Soc. 1976, 98, 2886–2894
[41] M. J. Kamlet, R. W. Taft, J. Am. Chem. Soc. 1976, 98, 377–383
[42] M. J. Kamlet, J. L. Abboud, R. W. Taft, J. Am. Chem. Soc. 1977, 99, 6027–6038
[43] C.-C. Chang, C.-J. Lin, ACM Trans. Intell. Syst. Technol. 2011, 2, 1–27
[44] D. Horvath, J. Brown, G. Marcou, A. Varnek, Challenges 2014, 5, 450
[45] A. Varnek, D. Fourches, D. Horvath, O. Klimchuk, C. Gaudin, P. Vayer, V. Solov´ev, F. Hoonakker, I. V. Tetko, G. Marcou, Curr. Comput.-Aided Drug Des. 2008, 4, 191–198
[46] M. Mathea, W. Klingspohn, K. Baumann, Mol. Inf. 2016, 35, 160–180